Durham University

Durham Research Online

Deposited in DRO:

21 January 2021

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Brettell, N. and Johnson, M. and Paesani, G. and Paulusma, D. (2020) 'Computing subset transversals in H-free graphs.', in WG 2020: Graph-Theoretic Concepts in Computer Science. , pp. 187-199. Lecture Notes in Computer Science., 12301

Further information on publisher's website:

https://doi.org/10.1007/978-3-030-60440-0_15

Publisher's copyright statement:

The final authenticated version is available online at https://doi.org/10.1007/978-3-030-60440-0_15

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full DRO policy for further details.

Durham University Library, Stockton Road, Durham DH1 3LY, United Kingdom
Tel : +44 (0)191 334 3042 | Fax : +44 (0)191 334 2971
https://dro.dur.ac.uk

Computing Subset Transversals in H-Free Graphs*

Nick Brettell**, Matthew Johnson, Giacomo Paesani, and Daniël Paulusma

Department of Computer Science, Durham University, UK
{nicholas.j.brettell,matthew.johnson2,giacomo.paesani,
daniel.paulusma}@durham.ac.uk

Abstract. We study the computational complexity of two well-known graph transversal problems, namely SUBSET FEEDBACK VERTEX SET and SUBSET ODD CYCLE TRANSVERSAL, by restricting the input to H-free graphs, that is, to graphs that do not contain some fixed graph H as an induced subgraph. By combining known and new results, we determine the computational complexity of both problems on H-free graphs for every graph H except when $H = sP_1 + P_4$ for some $s \geq 1$. As part of our approach, we introduce the SUBSET VERTEX COVER problem and prove that it is polynomial-time solvable for $(sP_1 + P_4)$-free graphs for every $s \geq 1$.

1 Introduction

The central question in Graph Modification is whether or not a graph G can be modified into a graph from a prescribed class $\mathcal{G}$ via at most k graph operations from a prescribed set S of permitted operations such as vertex or edge deletion. The *transversal* problems VERTEX COVER, FEEDBACK VERTEX SET and ODD CYCLE TRANSVERSAL are classical problems of this kind. For example, the VERTEX COVER problem is equivalent to asking if one can delete at most k vertices to turn G into a member of the class of edgeless graphs. The problems FEEDBACK VERTEX SET and ODD CYCLE TRANSVERSAL ask if a graph G can be turned into, respectively, a forest or a bipartite graph by deleting vertices.

We can relax the condition on belonging to a prescribed class to obtain some related *subset transversal* problems. We state these formally after some definitions. For a graph $G = (V, E)$ and a set $T \subseteq V$, an *(odd) T-cycle* is a cycle of G (with an odd number of vertices) that intersects T. A set $S_T \subseteq V$ is a *T-vertex cover*, a *T-feedback vertex set* or an *odd T-cycle transversal* of G if S_T has at least one vertex of, respectively, every edge incident to a vertex of T, every T-cycle, or every odd T-cycle. For example, let G be a star with centre vertex c, whose leaves form the set T. Then, both $\{c\} = V \setminus T$ and T are T-vertex covers of G but the first is considerably smaller than the second. See Figure 1 for some more examples.

* This paper received support from the Leverhulme Trust (RPG-2016-258).
** Corresponding Author: nicholas.j.brettell@durham.ac.uk

Subset Vertex Cover
Instance: a graph $G = (V, E)$, a subset $T \subseteq V$ and a positive integer k.
Question: does G have a T-vertex cover S_T with $|S_T| \leq k$?

Subset Feedback Vertex Set
Instance: a graph $G = (V, E)$, a subset $T \subseteq V$ and a positive integer k.
Question: does G have a T-feedback vertex set S_T with $|S_T| \leq k$?

Subset Odd Cycle Transversal
Instance: a graph $G = (V, E)$, a subset $T \subseteq V$ and a positive integer k.
Question: does G have an odd T-cycle transversal S_T with $|S_T| \leq k$?

Fig. 1. In both examples, the square vertices of the Petersen graph form a set T and the black vertices form an odd T-cycle transversal S_T, which is also a T-feedback vertex set. In the left example, $S_T \setminus T \neq \emptyset$, and in the right example, $S_T \subseteq T$.

The Subset Feedback Vertex Set and Subset Odd Cycle Transversal problems are well known. The Subset Vertex Cover problem is introduced in this paper, and we are not aware of past work on this problem. On general graphs, Subset Vertex Cover is polynomially equivalent to Vertex Cover: to solve Subset Vertex Cover remove edges in the input graph that are not incident to any vertex of T to yield an equivalent instance of Vertex Cover. However, this equivalence no longer holds for graph classes that are *not* closed under edge deletion.

As the three problems are NP-complete, we consider the restriction of the input to special graph classes in order to better understand which graph properties cause the computational hardness. Instead of classes closed under edge deletion, we focus on classes of graphs closed under vertex deletion. Such classes are called *hereditary*. The reasons for this choice are threefold. First, hereditary graph classes capture many well-studied graph classes. Second, every hereditary graph class $\mathcal{G}$ can be characterized by a (possibly infinite) set $\mathcal{F}_\mathcal{G}$ of forbidden induced subgraphs. This enables us to initiate a *systematic* study, starting from the case where $|\mathcal{F}_\mathcal{G}| = 1$. Third, we aim to extend and strengthen existing complexity results (that are for hereditary graph classes). If $\mathcal{F}_\mathcal{G} = \{H\}$ for some graph H, then $\mathcal{G}$ is *monogenic*, and every $G \in \mathcal{G}$ is *H-free*. Our research question is: *How does the structure of a graph H influence the computational complexity of a subset transversal problem for input graphs that are H-free?*

As a general strategy one might first try to prove that the restriction to H-free graphs is NP-complete if H contains a cycle or an induced claw (the 4-vertex star). This is usually done by showing, respectively, that the problem is NP-complete on graphs of arbitrarily large girth (the length of a shortest cycle) and on line graphs, which form a subclass of claw-free graphs. If this is the case, then it remains to consider the case where H has no cycle, and has no claw either. So H is a *linear forest*, that is, the disjoint union of one or more paths.

Existing Results As NP-completeness results for transversal problems carry over to subset transversal problems, we first discuss results on Feedback Vertex Set and Odd Cycle Transversal for H-free graphs. By Poljak's construction [33], Feedback Vertex Set is NP-complete for graphs of girth at least g for every integer $g \geq 3$. The same holds for Odd Cycle Transversal [8]. Moreover, Feedback Vertex Set [35] and Odd Cycle Transversal [8] are NP-complete for line graphs and thus for claw-free graphs. Hence, both problems are NP-complete for H-free graphs if H has a cycle or claw. Both problems are polynomial-time solvable for P_4-free graphs [5], for sP_2-free graphs for every $s \geq 1$ [8] and for $(sP_1 + P_3)$-free graphs for every $s \geq 1$ [13]. In addition, Odd Cycle Transversal is NP-complete for $(P_2 + P_5, P_6)$-free graphs [13]. Very recently, Abrishami et al. showed that Feedback Vertex Set is polynomial-time solvable for P_5-free graphs [1]. We summarize as follows ($F \subseteq_i G$ means that F is an induced subgraph of G; see Section 2 for the other notation used).

Theorem 1. *For a graph H,* Feedback Vertex Set *on H-free graphs is polynomial-time solvable if $H \subseteq_i P_5$, $H \subseteq_i sP_1 + P_3$ or $H \subseteq_i sP_2$ for some $s \geq 1$, and* NP*-complete if $H \supseteq_i C_r$ for some $r \geq 3$ or $H \supseteq_i K_{1,3}$.*

Theorem 2. *For a graph H,* Odd Cycle Transversal *on H-free graphs is polynomial-time solvable if $H = P_4$, $H \subseteq_i sP_1 + P_3$ or $H \subseteq_i sP_2$ for some $s \geq 1$, and* NP*-complete if $H \supseteq_i C_r$ for some $r \geq 3$, $H \supseteq_i K_{1,3}$, $H \supseteq_i P_6$ or $H \supseteq_i P_2 + P_5$.*

We note that no integer r is known such that Feedback Vertex Set is NP-complete for P_r-free graphs. This situation changes for Subset Feedback Vertex Set which is, unlike Feedback Vertex Set, NP-complete for split graphs (that is, $(2P_2, C_4, C_5)$-free graphs), as shown by Fomin et al. [16]. Papadopoulos and Tzimas [31,32] proved that Subset Feedback Vertex Set is polynomial-time solvable for sP_1-free graphs for any $s \geq 1$, co-bipartite graphs, interval graphs and permutation graphs, and thus P_4-free graphs. Some of these results were generalized by Bergougnoux et al. [2], who solved an open problem of Jaffke et al. [22] by giving an $n^{O(w^2)}$-time algorithm for Subset Feedback Vertex Set given a graph and a decomposition of this graph of mim-width w. This does not lead to new results for H-free graphs: a class of H-free graphs has bounded mim-width if and only if $H \subseteq_i P_4$ [7].

We are not aware of any results on Subset Odd Cycle Transversal for H-free graphs, but note that this problem generalizes Odd Multiway Cut, just as Subset Feedback Vertex Set generalizes Node Multiway Cut, another well-studied problem. We refer to [9,12,16,17,19,24,25,26,27,21] for further details, in particular for parameterized and exact algorithms for Subset Feedback Vertex Set and Subset Odd Cycle Transversal. These algorithms are beyond the scope of this paper.

Our Results By a significant extension of the known results for the two problems on H-free graphs we obtain two almost-complete dichotomies:

Theorem 3. *Let H be a graph with $H \neq sP_1 + P_4$ for all $s \geq 1$. Then* Subset Feedback Vertex Set *on H-free graphs is polynomial-time solvable if $H = P_4$ or $H \subseteq_i sP_1 + P_3$ for some $s \geq 1$ and* NP*-complete otherwise.*

Theorem 4. *Let H be a graph with $H \neq sP_1 + P_4$ for all $s \geq 1$. Then* Subset Odd Cycle Transversal *on H-free graphs is polynomial-time solvable if $H = P_4$ or $H \subseteq_i sP_1 + P_3$ for some $s \geq 1$ and* NP*-complete otherwise.*

Though the proved complexities of Subset Feedback Vertex Set and Subset Odd Cycle Transversal are the same on H-free graphs, the algorithm that we present for Subset Odd Cycle Transversal on $(sP_1 + P_3)$-free graphs is more technical compared to the algorithm for Subset Feedback Vertex Set, and considerably generalizes the transversal algorithms for (sP_1+P_3)-free graphs of [13]. There is further evidence that Subset Odd Cycle Transversal is a more challenging problem than Subset Feedback Vertex Set. For example, the best-known parameterized algorithm for Subset Feedback Vertex Set runs in $O^*(4^k)$ time [21], but the best-known run-time for Subset Odd Cycle Transversal is $O^*(2^{O(k^3 \log k)})$ [27]. Moreover, it is not known if there is an XP algorithm for Subset Odd Cycle Transversal in terms of mim-width in contrast to the known XP algorithm for Subset Feedback Vertex Set [2].

In Section 2 we introduce our terminology. In Section 3 we present some results for Subset Vertex Cover: the first result shows that Subset Vertex Cover is polynomial-time solvable for $(sP_1 + P_4)$-free graphs for every $s \geq 1$, and we later use this as a subroutine to obtain a polynomial-time algorithm for Subset Odd Cycle Transversal on P_4-free graphs. We present our results on Subset Feedback Vertex Set and Subset Odd Cycle Transversal in Sections 4 and 5, respectively. In Section 6 on future work we discuss Subset Vertex Cover in more detail.

2 Preliminaries

We consider undirected, finite graphs with no self-loops and no multiple edges. Let $G = (V, E)$ be a graph, and let $S \subseteq V$. The graph $G[S]$ is the subgraph of G induced by S. We write $G - S$ to denote the graph $G[V \setminus S]$. Recall that for a graph F, we write $F \subseteq_i G$ if F is an induced subgraph of G. The cycle and path on r vertices are denoted C_r and P_r, respectively. We say that S is *independent* if $G[S]$ is edgeless, and that S is a *clique* if $G[S]$ is *complete*, that is, contains every possible edge between two vertices. We let K_r denote the complete graph on r vertices, and sP_1 denote the graph whose vertices form an independent set of size s. A *(connected) component* of G is a maximal connected subgraph of G. The graph $\overline{G} = (V, \{uv \mid uv \notin E \text{ and } u \neq v\})$ is the *complement* of G. The *neighbourhood* of a vertex $u \in V$ is the set $N_G(u) = \{v \mid uv \in E\}$. For $U \subseteq V$, we let $N_G(U) = \bigcup_{u \in U} N(u) \setminus U$. The *closed* neighbourhoods of u and U are denoted by $N_G[u] = N_G(u) \cup \{u\}$ and $N_G[U] = N_G(U) \cup U$, respectively. We omit subscripts when there is no ambiguity.

Let $T \subseteq V$ be such that $S \cap T = \emptyset$. Then S is *complete* to T if every vertex of S is adjacent to every vertex of T, and S is *anti-complete* to T if there are no edges between S and T. In the first case, S is also said to be *complete* to $G[T]$, and in the second case we say it is *anti-complete* to $G[T]$.

We say that G is a *forest* if it has no cycles, and, furthermore, that G is a *linear forest* if it is the disjoint union of one or more paths. The graph G is *bipartite* if V can be partitioned into at most two independent sets. A graph is *complete bipartite* if its vertex set can be partitioned into two independent sets X and Y such that X is complete to Y. We denote such a graph by $K_{|X|,|Y|}$. If X or Y has size 1, the complete bipartite graph is a *star*; recall that $K_{1,3}$ is also called a claw. A graph G is a *split graph* if it has a bipartition (V_1, V_2) such that $G[V_1]$ is a clique and $G[V_2]$ is an independent set. A graph is split if and only if it is $(C_4, C_5, 2P_2)$-free [15].

Let G_1 and G_2 be two vertex-disjoint graphs. The *union* operation $+$ creates the disjoint union $G_1 + G_2$ of G_1 and G_2 (recall that $G_1 + G_2$ is the graph with vertex set $V(G_1) \cup V(G_2)$ and edge set $E(G_1) \cup E(G_2)$).

We also consider optimization versions of subset transversal problems, in which case we have instances (G, T) (instead of instances (G, T, k)). We say that a set $S \subseteq V(G)$ is a *solution* for an instance (G, T) if S is a T-transversal (of whichever kind we are concerned with). A solution S is *smaller* than a solution S' if $|S| < |S'|$, and a solution S is *minimum* if (G, T) does not have a solution smaller than S, and it is *maximum* if there is no larger solution. We will use the following general lemma, which was implicitly used in [32].

Lemma 1. *Let S be a minimum solution for an instance (G, T) of a subset transversal problem. Then $|S \setminus T| \leq |T \setminus S|$.*

Let $T \subseteq V$ be a vertex subset of a graph $G = (V, E)$. Recall that a cycle is a T-cycle if it contains a vertex of T. A subgraph of G is a *T-forest* if it has no T-cycles. Recall also that a cycle is odd if it has an odd number of edges. A subgraph of G is *T-bipartite* if it has no odd T-cycles. Recall that a set $S_T \subseteq V$ is a T-vertex cover, a T-feedback vertex set or an odd T-cycle transversal of G if S_T has at least one vertex of, respectively every edge incident to a vertex of T, every T-cycle, or every odd T-cycle. Note that S_T is a T-feedback vertex set if and only if $G[V \setminus S_T]$ is a T-forest, and S_T is an odd T-cycle transversal if and only if $G[V \setminus S_T]$ is T-bipartite. A *T-path* is a path that contains a vertex of T. A T-path is *odd* (or *even*) if the number of edges in the path is odd (or even, respectively).

We will use the following easy lemma, which proves that T-forests and T-bipartite graphs can be recognized in polynomial time. It combines results claimed but not proved in [27,32].

Lemma 2. *Let $G = (V, E)$ be a graph and $T \subseteq V$. Then deciding whether or not G is a T-forest or T-bipartite takes $O(n + m)$ time.*

3 Subset Vertex Cover

We need the following two results on Subset Vertex Cover (proofs omitted).

Lemma 3. Subset Vertex Cover *can be solved in polynomial time for P_4-free graphs.*

Lemma 4. *Let H be a graph. If* Subset Vertex Cover *is polynomial-time solvable for H-free graphs, then it is for $(P_1 + H)$-free graphs as well.*

Lemma 3, combined with s applications of Lemma 4, yields the following result.

Theorem 5. *For every integer $s \geq 1$,* Subset Vertex Cover *can be solved in polynomial time for $(sP_1 + P_4)$-free graphs.*

4 Subset Feedback Vertex Set

To prove Theorem 3. We require two lemmas. In the first lemma (whose proof we omit), the bound of $4s - 2$ is not necessarily tight, but suffices for our needs.

Lemma 5. *Let s be a non-negative integer, and let R be an $(sP_1 + P_3)$-free tree. Then either*

(i) *$|V(R)| \leq \max\{7, 4s - 2\}$, or*
(ii) *R has precisely one vertex r of degree more than 2 and at most $s - 1$ vertices of degree 2, each adjacent to r. Moreover, r has at least $3s - 1$ neighbours.*

We can extend "partial" solutions to full solutions in polynomial time as follows.

Lemma 6. *Let $G = (V, E)$ be a graph with a set $T \subseteq V$. Let $V' \subseteq V$ and $S'_T \subseteq V'$ such that S'_T is a T-feedback vertex set of $G[V']$, and let $Z = V \setminus V'$. Suppose that $G[Z]$ is P_3-free, and $|N_{G-S'_T}(Z)| \leq 1$. Then there is a polynomial-time algorithm that finds a minimum T-feedback vertex set S_T of G such that $S'_T \subseteq S_T$ and $V' \setminus S'_T \subseteq V \setminus S_T$.*

Proof. Since $G[Z]$ is P_3-free, it is a disjoint union of complete graphs. Let $G' = G - S'_T$, and consider a T-cycle C in G'. Then C contains at least one vertex of Z. If $N_{G'}(Z) = \emptyset$, then C is contained in a component of $G[Z]$. On the other hand, if $N_{G'}(Z) = \{y\}$, say, then y is a cut-vertex of G', so there exists a component $G[U]$ of $G[Z]$ such that C is contained in $G[U \cup \{y\}]$. Hence, we can consider each component of $G[Z]$ independently: for each component $G[U]$ it suffices to find the maximum subset U' of U such that $G[U' \cup N_{G'}(U)]$ contains no T-cycles. Then $U' \subseteq F_T$ and $U \setminus U' \subseteq S_T$, where $F_T = V \setminus S_T$.

Let $U \subseteq Z$ such that $G[U]$ is a component of $G[Z]$. Either $N_{G'}(U) \cap T = \emptyset$, or $N_{G'}(U) = \{y\}$ for some $y \in T$. First, consider the case where $N_{G'}(U) \cap T = \emptyset$. We find a set U' that is a maximum subset of U such that $G[U' \cup N_{G'}(U)]$ has no T-cycles. Clearly if $|U| = 1$, then we can set $U' = U$. If $|U'| \geq 3$, then, since U' is a clique, $U' \subseteq V \setminus T$. Thus, if $|U \setminus T| \geq 2$, then we set $U' = U \setminus T$. So it remains to consider when $|U| \geq 2$ but $|U \setminus T| \leq 1$. If there is some $u \in U$ that is anti-complete to $N_{G'}(U)$, then we can set U' to be any 2-element subset of U containing u. Otherwise $N_{G'}(U) = \{y\}$ and y is complete to U. In this case, for any $u \in U$, we set $U' = \{u\}$.

Now we may assume that $N_{G'}(U) = \{y\}$ and $y \in T$. Again, we find a set U' that is a maximum subset of U such that $G[U' \cup \{y\}]$ has no T-cycles. Partition U into $\{U_0, U_1\}$ where $u \in U_1$ if and only if u is a neighbour of y. Since $y \in V' \setminus S'_T$, observe that U' contains at most one vertex of U_1, otherwise $G[U' \cup \{y\}]$ has a T-cycle. Since U' is a clique, if $|U'| \geq 3$ then $U' \subseteq U \setminus T$. So if $|U_0 \setminus T| \geq 2$ and there is an element $u \in U_1 \setminus T$, then we can set $U' = \{u\} \cup (U_0 \setminus T)$. If $|U_0 \setminus T| \geq 2$ but $U_1 \setminus T = \emptyset$, then we can set $U' = U_0 \setminus T$. So we may now assume that $|U_0 \setminus T| \leq 1$. If $U_0 \neq \emptyset$ and $|U| \geq 2$, then we set U' to any 2-element subset of U containing some $u \in U_0$. Clearly if $|U| = 1$, then we can set $U' = U$. So it remains to consider when $U_0 = \emptyset$ and $|U_1| \geq 2$. In this case, we set $U' = \{u\}$ for an arbitrary $u \in U_1$. $\square$

We now prove the main result of this section.

Theorem 6. *For every integer $s \geq 0$,* Subset Feedback Vertex Set *can be solved in polynomial time for $(sP_1 + P_3)$-free graphs.*

Proof. Let $G = (V, E)$ be an $(sP_1 + P_3)$-free graph for some $s \geq 0$, and let $T \subseteq V$. We describe a polynomial-time algorithm for the optimization version of the problem on input (G, T). Let $S_T \subseteq V$ such that S_T is a minimum T-feedback vertex set of G, and let $F_T = V \setminus S_T$, so $G[F_T]$ is a maximum T-forest. Note that $G[F_T \cap T]$ is a forest. We consider three cases: either

1. $G[F_T \cap T]$ has at least $2s$ components;
2. $G[F_T \cap T]$ has fewer than $2s$ components, and each of these components consists of at most $\max\{7, 4s - 2\}$ vertices; or
3. $G[F_T \cap T]$ has fewer than $2s$ components, one of which consists of at least $\max\{8, 4s - 1\}$ vertices.

We describe polynomial-time subroutines that find a set F_T such that $G[F_T]$ is a maximum T-forest in each of these three cases, giving a minimum solution $S_T = V \setminus F_T$ in each case. We obtain an optimal solution by running each of these subroutines in turn: of the (at most) three potential solutions, we output the one with minimum size.

Case 1: $G[F_T \cap T]$ has at least $2s$ components.

We begin by proving a sequence of claims that describe properties of a maximum T-forest F_T, when in Case 1. Since G is (sP_1+P_3)-free, $F_T \cap T$ induces a P_3-free forest, so $G[F_T \cap T]$ is a disjoint union of graphs isomorphic to P_1 or P_2. Let $A \subseteq F_T \cap T$ such that $G[A]$ consists of precisely $2s$ components. Note that $|A| \leq 4s$. We also let $Y = N(A) \cap F_T$, and partition Y into $\{Y_1, Y_2\}$ where $y \in Y_1$ if y has only one neighbour in A, whereas $y \in Y_2$ if y has at least two neighbours in A.

Claim 1: $|Y_2| \leq 1$.

Proof of Claim 1. Let $v \in Y_2$. Then v has neighbours in at least $s+1$ of the components of $G[A]$, otherwise $G[A \cup \{v\}]$ contains an induced $sP_1 + P_3$. Note also that v has at most one neighbour in each component of $G[A]$, otherwise $G[F_T]$ has a T-cycle. Now suppose that Y_2 contains distinct vertices v_1 and v_2. Then, of the $2s$ components of $G[A]$, the vertices v_1 and v_2 each have some neighbour in $s+1$ of these components. So there are at least two components of $G[A]$ containing both a vertex adjacent to v_1, and a vertex adjacent to v_2. Let A' and A'' be the vertex sets of two such components. Then $A' \cup A'' \cup \{v_1, v_2\} \subseteq F_T$, but $G[A' \cup A'' \cup \{v_1, v_2\}]$ has a T-cycle; a contradiction. $\diamond$

Claim 2: $|Y| \leq 2s+1$.

Proof of Claim 2. By Claim 1, it suffices to prove that $|Y_1| \leq 2s$. We argue that each component of $G[A]$ has at most one neighbour in Y_1, implying that $|Y_1| \leq 2s$. Indeed, suppose that there is a component C_A of $G[A]$ having two neighbours in Y_1, say u_1 and u_2. Then $G[V(C_A) \cup \{u_1, u_2\}]$ contains an induced P_3 that is anti-complete to $A \setminus V(C_A)$, contradicting that G is (sP_1+P_3)-free. $\diamond$

Claim 3: Y_1 *is independent, and no component of* $G[A]$ *of size* 2 *has a neighbour in* Y_1*.*

Proof of Claim 3. Suppose that there are adjacent vertices u_1 and u_2 in Y_1. Let a_i be the unique neighbour of u_i in A for $i \in \{1,2\}$. Note that $a_1 \neq a_2$, for otherwise $G[F_T]$ has a T-cycle. Then $\{a_1, u_1, u_2\}$ induces a P_3, so $G[\{u_1, u_2\} \cup A]$ contains an induced $sP_1 + P_3$, which is a contradiction. We deduce that Y_1 is independent.

Now let $\{a_1, a_2\} \subseteq A$ such that $G[\{a_1, a_2\}]$ is a component of $G[A]$, and suppose that $u_1 \in Y_1$ is adjacent to a_1. Then a_1 is the unique neighbour of u_1 in A, so $G[\{u_1, a_1, a_2\}] \cong P_3$. Thus $G[\{u_1\} \cup A]$ contains an induced $sP_1 + P_3$, which is a contradiction. $\diamond$

Claim 4: Let $Z = V \setminus N[A]$*. Then* $N(Z) \cap F_T \subseteq Y_2$*.*

Proof of Claim 4. Suppose that there exists $y \in Y_1$ that is adjacent to a vertex $c \in Z$. Let a be the unique neighbour of y in A. Then $G[\{c, y\} \cup A]$ contains an induced $sP_1 + P_3$, which is a contradiction. So Y_1 is anti-complete to Z. Now, if $c \in Z$ is adjacent to a vertex in $N[A] \cap F_T$, then c is adjacent to y_2 where $Y_2 = \{y_2\}$. $\diamond$

We now describe the subroutine that finds an optimal solution in Case 1. In this case, for any maximum forest F_T, there exists some set $A \subseteq T$ of size at most $4s$ such that $A \subseteq F_T$, and $G[A]$ consists of exactly $2s$ components, each isomorphic

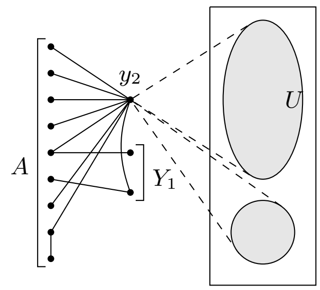

Fig. 2. An example of the structure obtained in Case 1 when $Y_2 = \{y_2\}$.

to either P_1 or P_2. Moreover, there is such an A for which $N(A) \cap T \subseteq S_T$. Thus we guess a set $A' \subseteq T$ in $O(n^{4s})$ time, discarding those sets that do not induce a forest with exactly $2s$ components, and those that induce a component consisting of more than two vertices.

For any such F_T and A', the set $N(A') \cap F_T$ has size at most $2s + 1$, by Claim 2. Thus, in $O(n^{2s+1})$ time, we guess $Y' \subseteq N(A')$ with $|Y'| \leq 2s + 1$, and assume that $Y' \subseteq F_T$ whereas $N(A') \setminus Y' \subseteq S_T$. Let Y_2' be the subset of Y' that contains vertices that have at least two neighbours in A'. We discard any sets Y' that do not satisfy Claims 1 or 3, or those sets for which $G[A' \cup Y']$ has a T-cycle on three vertices, one of which is the unique vertex of Y_2'.

Let $Z = V \setminus N[A']$ (for example, see Figure 2). Since $G[A']$ contains an induced sP_1, the subgraph $G[Z]$ is P_3-free. Now $N(Z) \cap F_T \subseteq Y_2'$ by Claim 4, where $|Y_2'| \leq 1$ by Claim 1. Thus, by Lemma 6, we can extend a partial solution $S_T' = N[A'] \setminus (A' \cup Y')$ of $G[N[A']]$ to a solution S_T of G, in polynomial time.

Case 2: $G[F_T \cap T]$ has at most $2s-1$ components, each of size at most $\max\{7, 4s-2\}$.

We guess sets $F \subseteq T$ and $S \subseteq V \setminus T$ such that $F_T \cap T = F$ and $S_T \setminus T = S$. Since F has size at most $(2s-1)\max\{7, 4s-2\}$ vertices, there are $O(n^{\max\{14s-7, 8s^2-8s+2\}})$ possibilities for F. By Lemma 1, we may assume that $|S_T \setminus T| \leq |F|$. So for each guessed F, there are at most $O(n^{\max\{14s-7, 8s^2-8s+2\}})$ possibilities for S. For each S and F, we set $S_T = (T \setminus F) \cup S$ and check, in $O(n+m)$-time by Lemma 2, if $G - S_T$ is a T-forest. In this way we exhaustively find all solutions satisfying Case 2, in $O(n^{\max^2\{14s-7, 8s^2-8s+2\}})$ time; we output the one of minimum size.

Case 3: $G[F_T \cap T]$ has at most $2s-1$ components, one of which has size at least $\max\{8, 4s-1\}$.

By Lemma 5, there is some subset $B_T \subseteq F_T \cap T$ such that $|B| \geq \max\{8, 4s-1\}$, and $G[B]$ is a component of $G[F_T \cap T]$ that is a tree satisfying Lemma 5(ii). In particular, there is a unique vertex $r \in B$ such that r has degree more than 2 in $G[B]$. Moreover, $G[F_T]$ has a component $G[D]$ that contains B, where $G[D]$ is a tree that also satisfies Lemma 5(ii). Note that there are at most $s-1$ vertices in $N_{G[B]}(r)$ having a neighbour in $V \setminus T$.

We guess a set $B' \subseteq T$ such that $|B'| = \max\{8, 4s-1\}$. We also guess a set $L' \subseteq V \setminus T$ such that $|L'| \leq s-1$. Let $D' = B' \cup L'$. We check that $G[D']$ has the following properties:

- $G[D']$ is a tree,
- $G[D']$ has a unique vertex r' of degree more than 2, with $r' \in B'$,
- $G[D']$ has at most $s-1$ vertices with distance 2 from r', and each of these vertices has degree 1, and
- each vertex $v \in L'$ has degree 1 in $G[D']$, and distance 2 from r'.

We assume that D' induces a subtree of the large component $G[D]$, where $r = r'$, and D' contains r, all neighbours of r with degree 2 in $G[D]$, and all vertices at distance 2 from r. In other words, $G[D']$ can be obtained from $G[D]$ by deleting some subset of the leaves of $G[D]$ that are adjacent to r. In particular, $D' \subseteq F_T$. We also assume that L' is the set of all vertices of $V(D) \setminus T$ that have distance 2 from r.

It follows from these assumptions that $N(D' \setminus \{r\}) \setminus \{r\} \subseteq S_T$. Let $Z = V \setminus N[D' \setminus \{r\}]$, and observe that each $z \in Z$ has at most one neighbour in D' (if it has such a neighbour, this neighbour is r). So $N(Z) \cap F_T \subseteq \{r\}$.

Towards an application of Lemma 6, we claim that $G[Z]$ is P_3-free. Let $B_1 = B' \cap N(r)$. As r has at least $3s-1$ neighbours in $G[B']$, by Lemma 5, $G[B_1]$ contains an induced sP_1. Moreover, $N(B_1) \cap F_T \subseteq D'$. Since G is (sP_1+P_3)-free, $G[Z]$ is P_3-free. We now apply Lemma 6, which completes the proof. □

We are now ready to prove Theorem 3.

Theorem 3 (restated). *Let H be a graph with $H \neq sP_1+P_4$ for all $s \geq 1$. Then* Subset Feedback Vertex Set *on H-free graphs is polynomial-time solvable if $H = P_4$ or $H \subseteq_i sP_1 + P_3$ for some $s \geq 1$ and is* NP*-complete otherwise.*

Proof. If H has a cycle or claw, we use Theorem 1. The cases $H = P_4$ and $H = 2P_2$ follow from the corresponding results for permutation graphs [31] and split graphs [16]. The remaining case $H \subseteq_i sP_1+P_3$ follows from Theorem 6. □

5 Subset Odd Cycle Transversal

At the end of this section we prove Theorem 4. We show three new results (proofs omitted). Our first result uses the reduction of [31] which proved the analogous result for Subset Feedback Vertex Set. Our third result is the main result of this section. Its proof uses the same approach as the proof of Theorem 6 but we need more advanced arguments for distinguishing cycles according to parity.

Theorem 7. Subset Odd Cycle Transversal *is* NP*-complete for the class of split graphs (or equivalently, $(C_4, C_5, 2P_2)$-free graphs).*

Theorem 8. Subset Odd Cycle Transversal *can be solved in polynomial time for P_4-free graphs.*

Theorem 9. *For every integer $s \geq 0$,* Subset Odd Cycle Transversal *can be solved in polynomial time for $(sP_1 + P_3)$-free graphs.*

We are now ready to prove our almost-complete classification.

Theorem 4 (restated). *Let H be a graph with $H \neq sP_1 + P_4$ for all $s \geq 1$. Then* Subset Odd Cycle Transversal *on H-free graphs is polynomial-time solvable if $H = P_4$ or $H \subseteq_i sP_1 + P_3$ for some $s \geq 1$ and* NP*-complete otherwise.*

Proof. If H has a cycle or claw, we use Theorem 2. The cases $H = P_4$ and $H = 2P_2$ follow from Theorems 7 and 8, respectively. The remaining case, where $H \subseteq_i sP_1 + P_3$, follows from Theorem 9. □

6 Conclusions

We gave almost-complete classifications of the complexity of Subset Feedback Vertex Set and Subset Odd Cycle Transversal for H-free graphs. The only open case in each classification is when $H = sP_1 + P_4$ for some $s \geq 1$, which is also open for Feedback Vertex Set and Odd Cycle Transversal for H-free graphs. Our proof techniques for $H = sP_1 + P_3$ do not carry over and new structural insights are needed in order to solve the missing cases where $H = sP_1 + P_4$ for $s \geq 1$.

We also introduced the Subset Vertex Cover problem and showed that this problem is polynomial-time solvable on $(sP_1 + P_4)$-free graphs for every $s \geq 0$. Lokshtanov et al. [28] proved that Vertex Cover is polynomial-time solvable for P_5-free graphs. Grzesik et al. [18] extended this result to P_6-free graphs. What is the complexity of Subset Vertex Cover for P_5-free graphs? Does there exist an integer $r \geq 5$ such that Subset Vertex Cover is NP-complete for P_r-free graphs. By Poljak's construction [33], Vertex Cover is NP-complete for H-free graphs if H has a cycle. However, Vertex Cover becomes polynomial-time solvable on $K_{1,3}$-free graphs [29,34]. We did not research the complexity of Subset Vertex Cover on $K_{1,3}$-free graphs and also leave this as an open problem for future work.

Finally, several related transversal problems have been studied but not yet for H-free graphs. For example, the parameterized complexity of Even Cycle Transversal and Subset Even Cycle Transversal has been addressed in [30] and [24], respectively. Moreover, several other transversal problems have been studied for H-free graphs, but not the subset version: for example, Connected Vertex Cover, Connected Feedback Vertex Set and Connected Odd Cycle Transversal, and also for Independent Feedback Vertex Set and Independent Odd Cycle Transversal; see [4,8,14,23] for a number of recent results. It would be interesting to solve the subset versions of these transversal problems for H-free graphs and to determine the connections amongst all these problems in a more general framework.

References

1. T. Abrishami, M. Chudnovsky, M. Pilipczuk, P. Rzążewski, and P. Seymour. Induced subgraphs of bounded treewidth and the container method. *CoRR*, abs/2003.05185, 2020.
2. B. Bergougnoux, C. Papadopoulos, and J. A. Telle. Node multiway cut and subset feedback vertex set on graphs of bounded mim-width. *Proc. WG 2020, LNCS*, to appear.
3. H. L. Bodlaender and R. H. Möhring. The pathwidth and treewidth of cographs. *SIAM Journal on Discrete Mathematics*, 6:181–188, 1993.
4. M. Bonamy, K. K. Dabrowski, C. Feghali, M. Johnson, and D. Paulusma. Independent feedback vertex set for P_5-free graphs. *Algorithmica*, 81:1342–1369, 2019.
5. A. Brandstädt and D. Kratsch. On the restriction of some NP-complete graph problems to permutation graphs. *Proc. FCT 1985, LNCS*, 199:53–62, 1985.
6. A. Brandstädt, V. B. Le, and J. P. Spinrad. *Graph Classes: A Survey*, volume 3 of *SIAM Monographs on Discrete Mathematics and Applications*. SIAM, 1999.
7. N. Brettell, J. Horsfield, A. Munaro, G. Paesani, and D. Paulusma. Bounding the mim-width of hereditary graph classes. *CoRR*, abs/2004.05018, 2020.
8. N. Chiarelli, T. R. Hartinger, M. Johnson, M. Milanič, and D. Paulusma. Minimum connected transversals in graphs: New hardness results and tractable cases using the price of connectivity. *Theoretical Computer Science*, 705:75–83, 2018.
9. R. Chitnis, F. V. Fomin, D. Lokshtanov, P. Misra, M. S. Ramanujan, and S. Saurabh. Faster exact algorithms for some terminal set problems. *Journal of Computer and System Sciences*, 88:195–207, 2017.
10. D. G. Corneil, H. Lerchs, and L. S. Burlingham. Complement reducible graphs. *Discrete Applied Mathematics*, 3:163–174, 1981.
11. D. G. Corneil, Y. Perl, and L. K. Stewart. A linear recognition algorithm for cographs. *SIAM Journal on Computing*, 14:926–934, 1985.
12. M. Cygan, M. Pilipczuk, M. Pilipczuk, and J. O. Wojtaszczyk. Subset feedback vertex set is fixed-parameter tractable. *SIAM Journal on Discrete Mathematics*, 27:290–309, 2013.
13. K. K. Dabrowski, C. Feghali, M. Johnson, G. Paesani, D. Paulusma, and P. Rzążewski. On cycle transversals and their connected variants in the absence of a small linear forest. *Algorithmica*, to appear.
14. K. K. Dabrowski, M. Johnson, G. Paesani, D. Paulusma, and V. Zamaraev. On the price of independence for vertex cover, feedback vertex set and odd cycle transversal. *Proc. MFCS 2018, LIPIcs*, 117:63:1–63:15, 2018.
15. S. Földes and P. L. Hammer. Split graphs. *Congressus Numerantium*, 19:311–315, 1977.
16. F. V. Fomin, P. Heggernes, D. Kratsch, C. Papadopoulos, and Y. Villanger. Enumerating minimal subset feedback vertex sets. *Algorithmica*, 69:216–231, 2014.
17. P. A. Golovach, P. Heggernes, D. Kratsch, and R. Saei. Subset feedback vertex sets in chordal graphs. *Journal of Discrete Algorithms*, 26:7–15, 2014.
18. A. Grzesik, T. Klimošová, M. Pilipczuk, and M. Pilipczuk. Polynomial-time algorithm for maximum weight independent set on P_6-free graphs. *Proc. SODA 2019*, pages 1257–1271, 2019.
19. E. C. Hols and S. Kratsch. A randomized polynomial kernel for subset feedback vertex set. *Theory of Computing Systems*, 62(1):63–92, 2018.
20. J. E. Hopcroft and R. E. Tarjan. Algorithm 447: Efficient algorithms for graph manipulation. *Communications of the ACM*, 16(6):372–378, 1973.

21. Y. Iwata, M. Wahlström, and Y. Yoshida. Half-integrality, lp-branching, and FPT algorithms. *SIAM Journal on Computing*, 45(4):1377–1411, 2016.
22. L. Jaffke, O. Kwon, and J. A. Telle. Mim-width II. the feedback vertex set problem. *Algorithmica*, 82(1):118–145, 2020.
23. M. Johnson, G. Paesani, and D. Paulusma. Connected vertex cover for $(sP_1 + P_5)$-free graphs. *Algorithmica*, 82:20–40, 2020.
24. N. Kakimura, K. Kawarabayashi, and Y. Kobayashi. Erdős-Pósa property and its algorithmic applications: parity constraints, subset feedback set, and subset packing. *Proc. SODA 2012*, pages 1726–1736, 2012.
25. K. Kawarabayashi and Y. Kobayashi. Fixed-parameter tractability for the subset feedback set problem and the S-cycle packing problem. *Journal of Combinatorial Theory, Series B*, 102(4):1020–1034, 2012.
26. S. Kratsch and M. Wahlström. Representative sets and irrelevant vertices: New tools for kernelization. *Proc. FOCS 2012*, pages 450–459, 2012.
27. D. Lokshtanov, P. Misra, M. S. Ramanujan, and S. Saurabh. Hitting selected (odd) cycles. *SIAM Journal on Discrete Mathematics*, 31(3):1581–1615, 2017.
28. D. Lokshtanov, M. Vatshelle, and Y. Villanger. Independent set in P_5-free graphs in polynomial time. *Proc. SODA 2014*, pages 570–581, 2014.
29. G. J. Minty. On maximal independent sets of vertices in claw-free graphs. *Journal of Combinatorial Theory, Series B*, 28(3):284–304, 1980.
30. P. Misra, V. Raman, M. S. Ramanujan, and S. Saurabh. Parameterized algorithms for even cycle transversal. *Proc. WG 2012*, pages 172–183, 2012.
31. C. Papadopoulos and S. Tzimas. Polynomial-time algorithms for the subset feedback vertex set problem on interval graphs and permutation graphs. *Discrete Applied Mathematics*, 258:204–221, 2019.
32. C. Papadopoulos and S. Tzimas. Subset feedback vertex set on graphs of bounded independent set size. *Theoretical Computer Science*, 814:177–188, 2020.
33. S. Poljak. A note on stable sets and colorings of graphs. *Commentationes Mathematicae Universitatis Carolinae*, 15:307–309, 1974.
34. N. Sbihi. Algorithme de recherche d'un stable de cardinalité maximum dans un graphe sans étoile. *Discrete Mathematics*, 29(1):53–76, 1980.
35. E. Speckenmeyer. *Untersuchungen zum Feedback Vertex Set Problem in ungerichteten Graphen*. PhD thesis, Universität Paderborn, 1983.